**Raid技术拓展研究**

**目录**

[RAID技术简介 1](#_Toc13055011)

[RAID 0 2](#_Toc13055012)

[RAID 1 4](#_Toc13055013)

[RAID 5 6](#_Toc13055014)

[RAID 10 8](#_Toc13055015)

[RAID 01和RAID 10的选择问题 9](#_Toc13055016)

[工作方式： 10](#_Toc13055017)

[安全性： 10](#_Toc13055018)

[读写速度： 11](#_Toc13055019)

[结论： 11](#_Toc13055020)

[RAID 5和RAID 10的差异 11](#_Toc13055021)

[空间利用率： 12](#_Toc13055022)

[安全性： 12](#_Toc13055023)

[Cache： 12](#_Toc13055024)

[读操作的差异： 13](#_Toc13055025)

[连续写的差异： 14](#_Toc13055026)

[离散写的差异： 14](#_Toc13055027)

[总结： 15](#_Toc13055028)

## RAID技术简介

RAID(独立冗余磁盘阵列) 技术是为了组合小的廉价磁盘来代替大的昂贵磁盘, 同时希望磁盘失效时不会使对数据的访问受损失而开发出一定水平的数据保护技术。RAID可以充分发挥出多块硬盘的优势，可以提升硬盘速度，增大容量，提供容错功能够确保数据安全性，易于管理的优点，在任何一块硬盘出现问题的情况下都可以继续工作，不会受到损坏硬盘的影响。

RAID技术的分类有RAID 0 , RAID 1 , RAID 2, RAID 3 ,RAID 4 , RAID 5 , RAID 6 , RAID 10 , RAID 01 , RAID 50等。RAID 10和RAID 01的工作方式有些类似，但由于RAID 10的安全性要比RAID 01好得多，在日常应用中RAID 10的使用次数要比RAID 01多。

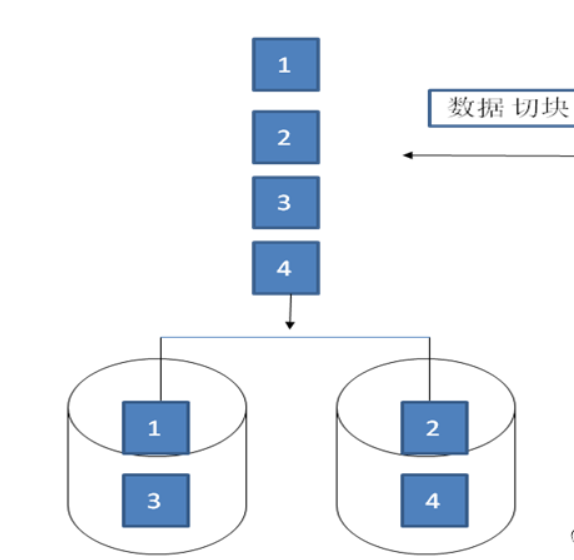
**RAID 0，RAID 1，RAID 5，RAID 10这四种比较典型，以下主要讨论这四种RAID技术以及比较他们性能的差距。**

## ****RAID 0****

RAID 0亦称为带区集。一般情况下它将两个以上的硬盘串联起来，连成一个容量更大的硬盘群。

**工作方式：**

先把硬盘分切出等量的区块，当文件要写入硬盘中时，把数据依据硬盘区块大小切割好，再依序交错存入硬盘。如两块物理磁盘应用了RAID 0技术，若一个数据文件有1GB，则每个磁盘会各存入500MB，读取数据则反之。每块硬盘只要负担一半的数据传输任务。如下图：



**优点：**

在存放数据时，分段后分散存储在这些磁盘中，因为读写时都可以并行处理，所以在所有的级别中，RAID 0的读性能高，连续写性能和随机写性能也高，速度是最快的。

低成本，把廉价的容量小的磁盘，在逻辑上组成一个大容量的磁盘，逻辑磁盘容量为所有物理磁盘容量之和。

**缺点：**

提高磁盘的性能和吞吐量但是RAID 0既没有冗余功能，也不具备容错能力，如果一个磁盘（物理）损坏，所有数据都会丢失。

**应用：**

应用RAID 0技术的硬盘，容量一定要相同。

用n块硬盘选择合理的带区大小创建带区集，最好是为每一块硬盘都配备一个专门的磁盘控制器,在电脑数据读写时同时向n块磁盘读写数据,速度提升n倍。提高系统的性能。RAID 0特别适用于对性能要求较高，而对数据安全性要求不高的领域，如图形工作站、个人用户、PC服务器等。

## RAID 1

RAID 1称为磁盘镜像，他把一个磁盘的数据镜像到另一个磁盘上，在不影响性能情况下最大限度的保证系统的可靠性和可修复性上，具有很高的数据冗余能力，但磁盘利用率为50%。

**工作方式：**

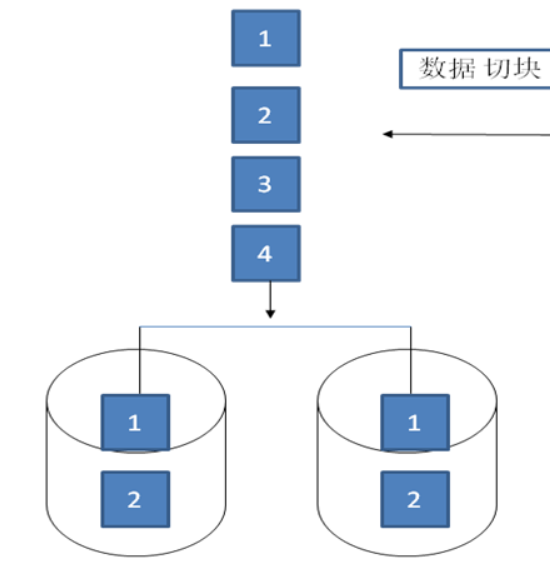
**写：**

先把硬盘分切出等量的区块，当文件要写入磁盘中时，把数据依据磁盘区块大小切割好，再存入各磁盘各一份，结果是把用户写入硬盘的数据百分之百地自动复制到另外一个硬盘上。

**读：**

系统先从RAID 1的源盘读取数据，如果读取数据成功，则系统不去管备份盘上的数据；如果读取源盘数据失败，则系统转而读取备份盘上的数据。

如下图：



**优点：**

RAID 1的每一个磁盘都具有一个对应的镜像盘，任何时候数据都同步镜像，系统可以从一组镜像盘中的任何一个磁盘读取数据。

只要系统中任何一对镜像盘中至少有一块磁盘可以使用，甚至可以在一半数量的硬盘出现问题时系统都可以正常运行。

具有很高的数据冗余能力，安全性最强，可靠性最高。

如果用两个不同大小的磁盘建立RAID 1，实际可用空间为较小的那个磁盘，较大的磁盘多出来的空间也可以分区成一个区来使用，不会造成浪费。

在一些多线程操作系统中能有很好的读取速度。

RAID 1磁盘阵列一般支持“热交换”，就是说阵列中硬盘的移除或替换可以在系统运行时进行，无须中断退出系统。更换完毕后只需要从正常的磁盘恢复数据。

**缺点：**

是所有RAID中磁盘利用率最低的一个级别，磁盘所能使用的空间只有磁盘容量总和的一半，系统成本高。

磁盘控制器的负载相当大，写入速度较慢。

更换新盘后原有数据会需要很长时间同步镜像，外界对数据的访问不会受到影响，但是这时整个系统的性能有所下降。

**应用：**

至少需要两块磁盘，磁盘容量不必相同。

尤其适用于存放重要数据，如服务器和数据库存储等领域。

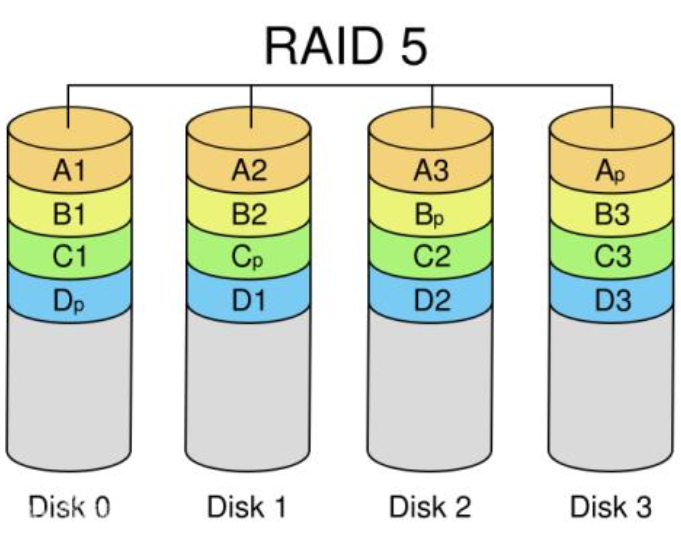
要及时地更换损坏的硬盘并利用备份数据重新建立Mirror，避免备份盘在发生损坏时，造成不可挽回的数据损失。

## RAID 5

RAID 5是一种储存性能、数据安全和存储成本兼顾的存储解决方案。RAID 5不是对存储的数据进行备份，而是把数据和相对应的奇偶校验信息存储到组成RAID5的各个磁盘上，并且奇偶校验信息和相对应的数据分别存储于不同的磁盘上。当RAID5的一个磁盘数据发生损坏后，RAID会自动利用剩下的数据和相应的奇偶校验信息去恢复被损坏的数据。

**工作方式：**

对数据进行分段，并做异或操作，各个硬盘在SCSI控制卡的控制下平行动作，存数据的同时存异或的结果，校验数据以循环的方式放在磁盘。如下图，Ap为A1、A2、A3的异或结果。



即当Disk 0保存的数据为1，Disk 1保存的数据为1，Disk 2保存的数据为0时，Disk 3中的数据以1、1和0作异或运算即为0。数据存储时经过分块。

**优点：**

RAID 5具有和RAID 0相近似的数据读取速度，但是因为要处理奇偶校验，故比RAID 0稍微慢些；RAID 5可以为系统提供数据安全保障，但保障程度要比RAID 1低，而由于多个数据对应一个奇偶校验，信息磁盘空间利用率要比RAID 1高，成本也相对较低。

支持热备盘，保障数据的安全。

**缺点：**

只允许单盘故障，一盘出现故障必须尽快处理。

有盘坏情况下，RAID 5 IO/CPU性能急剧下降。

**应用：**

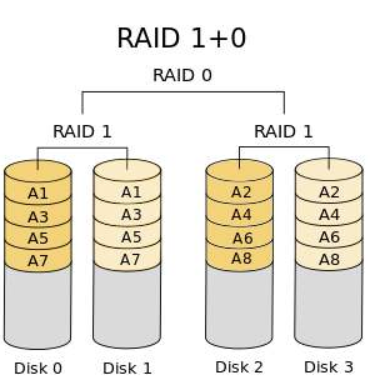
至少需要三块磁盘。可用的容量：n-1/n(n为磁盘数且n>=3)。

适合磁盘数量不多，且对数据安全有一定要求，多读少写的情况，如银行、金融、股市、数据库等大型数据处理中心OLTP应用。

## RAID 10

RAID 10也被称为镜像阵列条带。RAID 10读取数据是跨磁盘读取，类似于RAID 0；每个磁盘有镜像磁盘，即数据百分百冗余，类似RAID 1。RAID 10利用了RAID 0极高的读写效率和RAID 1较高的数据保护、恢复能力。

**工作方式：**



如上图，数据先进行镜像操作，再分区数据，再将所有硬盘分成两组，视为RAID 0的组合，进行存储。最后数据被分别存入Disk 0、Disk 2，而备份被存到Disk 1和Disk 3。

**优点：**

由于使用RAID 0作为执行等级，因此具有较高的I/O宽带，读写速度快。

以RAID 1为数据保护阵列，它具有与RAID 1一样的容错能力，用于容错处理的[系统开销](https://baike.baidu.com/item/%E7%B3%BB%E7%BB%9F%E5%BC%80%E9%94%80)与单独的[镜像](https://baike.baidu.com/item/%E9%95%9C%E5%83%8F)操作基本一样。

只有一块盘出错时，只需要更换新的硬盘，并应用RAID 10来进行数据恢复，恢复数据过程中系统仍能正常工作，维护难度不大。

容易实现，控制器无需强劲处理能力，通常也无需大缓存。

RAID卡支持热备盘的情况下，安装热备盘安全系数更高。

**缺点：**

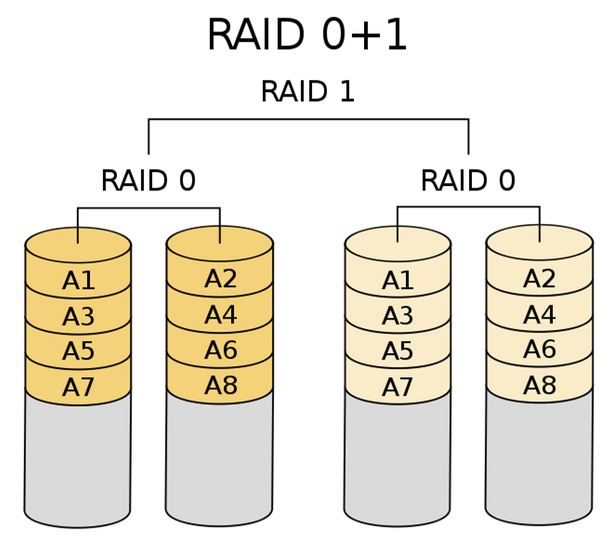
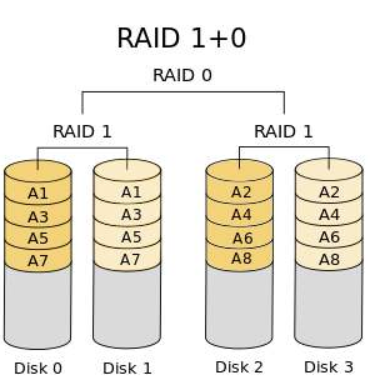
对磁盘数量要求高，而且只能使用其中一半的磁盘容量，如果磁盘大小不一，使用率更低。磁盘利用率低，故成本高。

源磁盘和镜像磁盘同时出错，数据不可恢复。

**应用：**

RAID 10需要4 + 2n个磁盘（n>=0)。RAID 10的高安全性、高读写速度结构适用于数据库存储服务器、校园FTP服务器等需要高性能、高容错但对容量要求不大的场合。

## RAID 01和RAID 10的选择问题



## 工作方式：

在硬件上，RAID 01和RAID 10 要求的最少硬盘数一样是4，而在具体的工作方式上RAID 01与RAID 10略有不同。

以上图为例，使用4个硬盘。RAID 01先将所有的硬盘分为两组，变成RAID 1的最低组合，可见Disk1、Disk2为一组作为源磁盘，Disk3、Disk4为一组作为备份盘。再将两组硬盘各自视为RAID 0运作，即源磁盘组中分段存储数据，而备份盘组中按照源磁盘分段备份数据。RAID 10则是先镜射再分区数据。

## 安全性：

在实际应用中，4个磁盘组成RAID 01（如上图），假设Disk0故障离线，那么作为源磁盘组的RAID 0组整组离线（左边RAID 0组离线）。此时可使用的硬盘只剩下右边Disk2、Disk3，RAID 01的设计退化成了RAID 0，磁盘可以继续以RAID 0工作至Disk2、Disk3任一个离线。不再有冗余磁盘意味着数据安全无保障。实际上RAID 01只允许一个磁盘故障，安全性并不理想。

实际中以4个硬盘实现RAID 10为例（可参考上图），假设Disk0故障离线，Disk1并不会同时离线，RAID 1允许一个盘离线。从整个RAID 10结构上看，结构仍然可使用。左边的RAID 1失去镜像盘，安全性降低；右边RAID 1仍然保证高安全级别。因损坏两块磁盘而导致整个RAID 10崩溃的几率只有2/（n-1）（n为硬盘数量）。

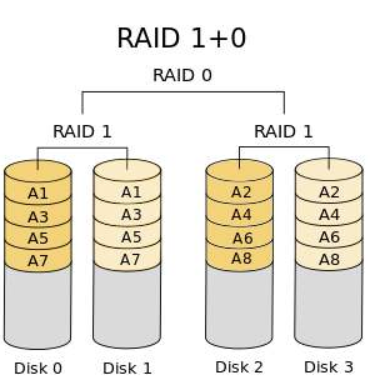
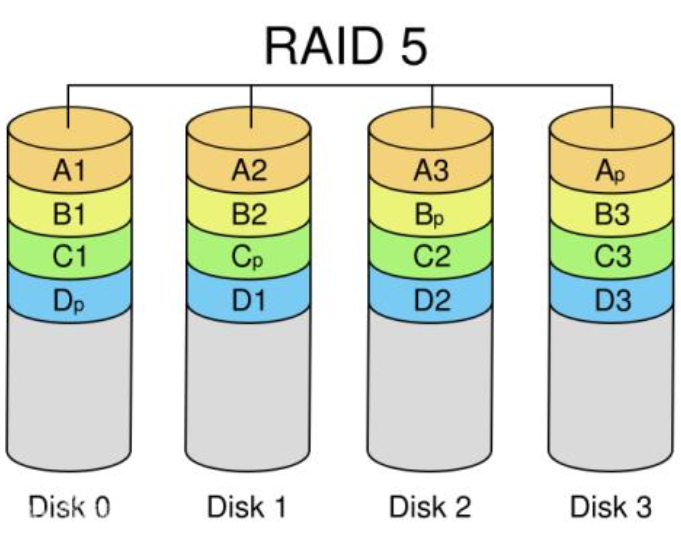
## 读写速度：

由于RAID 01和RAID 10都是基于RAID 0和RAID 1，读写速度相同，均为n/2（n为硬盘数）的理论宽带。

## 结论：

综上所述，RAID 10的可用性比RAID 01高得多。RAID 01虽说有一定的保障数据安全的能力，但是经过实践的检验，RAID 01出现故障概率比RAID 10大得多，这就说明RAID 01是华而不实的结构。实际应用RAID 10技术非常多，而适用RAID 01技术的场合非常少。

## RAID 5和RAID 10的差异



以下以RAID5选择3D+1P的RAID方案，RAID10选择2D+2D的RAID方案作讨论。

## 空间利用率：

RAID 5只有25%的空间用作冗余，故实际硬盘使用率为75%；RAID 10有50%的磁盘空间用来做冗余，实际上只有50%的磁盘空间可使用。由此可见RAID 5的磁盘使用率比RAID 10高。

## 安全性：

以上图为例，假设RAID 5的Disk0故障离线，此时RAID 5仍能正常工作，但是由于需要进行检验还原数据，系统CPU负载加重，系统总体性能会下降。缺少一个磁盘的RAID 5已经不具有保护数据安全的能力，如果有磁盘再故障离线，整个RAID 5将崩溃，数据丢失。

RAID 5支持热备盘，自动恢复数据，这也是提高RAID 5安全性的手段。

假设RAID 10的Disk0故障离线，RAID 10正常运行，并且系统I/O、CPU工作正常。只要Disk1正常工作，就算Disk2或Disk3任意一个离线系统还能正常工作，只不过在安全级别上RAID 10变成了RAID 0。

总的来说，RAID 5最多只允许一个磁盘离线，而RAID 10最多允许两个磁盘离线，RAID 10的安全性比RAID 5高。

## Cache：

Cache（高速缓冲存储器）具有相当高的存储速度，用来暂时存储数据,能使CPU与内存更协调合作。作为高端存储，cache已经是整个存储的核心所在，就是中低端存储，也有很大的cache存在，包括最简单的RAID卡，一般都包含有几十，甚至几百兆的RAID cache。一般阵列只要求数据写到cache就算完成了写操作，当cache的数据积累到一定程度，阵列才把数据刷到磁盘，可以实现批量的写入。Cache也是具有冗余的，一般是通过镜像与UPS。

Cache在读数据方面的作用一样不可忽视，因为如果所需要读取的数据能在cache中命中的话，将大大减少磁盘寻道所需要的时间。因为磁盘从开始寻道到找到数据，一般都在6ms以上，而这个时间，对于那些密集型I/O的应用可能不是太理想。但是，如果能在cache保存的数据中命中，一般响应时间则可以缩短在1ms以内。

由此可见，RAID卡配合cache在全局上实现数据存储大大加快了读写的速度，但是实际上读取数据在cache中的命中率并不是百分百。

## 读操作的差异：

磁盘阵列读操作关键在于cache的命中率，在有RAID卡建立RAID并且RAID 5和RAID 10都无单盘故障的情况下，他们读取数据的速度都是相当的；在都存在单盘故障并且有RAID卡建立RAID的情况下，由于RAID 5需要做校验，但是考虑到数据可能在cache中命中，从整体的角度看RAID 5会比RAID 10慢一些；在无RAID卡且单盘故障的情况下，RAID 5的读速度跌得非常厉害，而RAID 10不受影响。

## 连续写的差异：

连续写的过程一般表示写入连续的大批量的数据，如媒体数据流，很大的文件等。RAID 5校验是在cache中完成，如4块盘的RAID5，可以先在内存中计算好校验，同时写入3个数据+1个校验。而RAID10只能同时写入2个数据+2个镜相。

在连续写操作过程，如果有写cache存在，并且算法没有问题的话，RAID5比RAID10甚至会更好一些，如果不是连续性的强连续写，只要不达到磁盘的写极限，差别并不是太大。（这里要假定存储有一定大小足够的写cache，而且计算校验的cpu不会出现瓶颈）。

## 离散写的差异：

绝大部分数据库应用，如ERP系统等等在数据写入的时候其实都是离散写。Oracle 数据库每次写一个数据块的数据，如4K。由于每次写入的量不是很大，而且写入的次数非常频繁，因此联机日志看起来会像是连续写。对RAID 5来说，由于不保证能够添满RAID 5的一个条带（保证每张盘都能写入），所以在写入已存在数据的条带中，RAID 5很多时候更加偏向于离散写入。

在写入已存在数据的条带中时，RAID 5先要读出数据，接着读出校验，在cache中重新计算完校验后，才把新数据和新校验写入磁盘；RAID 10只需要把新数据直接写入。由此可见RAID 5需要4个IO,为RAID 10只需要2个IO。

如果RAID 5能在cache中命中数据，那么离散写就不需要4个IO。可见RAID 5对cache有非常强的依赖性。当然cache也能提升RAID 10的写速度，只是RAID 10对cache的依赖程度没有RAID 5高。

## 总结：

根据RAID 5和RAID 10各自的特性得知，RAID 5的空间利用率和高强度读写快的特性，使得RAID 5尤其适用在对数据安全要求不是极致，存储空间要求很大并且连续读写强度高的系统上，如媒体存储仓库。

RAID 10在离散数据写上的优势，以及其超高的安全性，使得它非常适合应用在银行数据库、股票交易数据库、普通小数据量频繁写入系统等。